Entangled qutrits violate local realism stronger than qubits –an analytical proof

Jing-Ling Chen,[1,3] Dagomir Kaszlikowski,[1,2] L. C. Kwek,[1,4] Marek Żukowski,[5] and C. H. Oh[1]
[1] *Department of Physics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260*
[2] *Instytut Fizyki Doświadczalnej, Uniwersytet Gdański, PL-80-952, Gdańsk, Poland,*
[3] *Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, P.O. Box 8009(26), Beijing 100088, People's Republic of China*
[4] *National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 639798*
[5] *Instytut Fizyki Teoretycznej i Astrofizyki, Uniwersytet Gdański, PL-80-952, Gdańsk, Poland.*

In Kaszlikowski *et al.* [Phys. Rev. Lett. **85**, 4418 (2000)], it has been shown numerically that the violation of local realism for two maximally entangled N-dimensional ($3 \leq N$) quantum objects is stronger than for two maximally entangled qubits and grows with N. In this paper we present the analytical proof of this fact for $N = 3$.

Since the formulation of Bell theorem [1], various forms of so called Bell inequalities, see for instance [2], have been devised to investigate the possibility (or lack of such possibility) of local realistic description of correlations observed in various quantum systems such as M entangled N dimensional quantum objects. The main advantage of this approach is its simplicity. The drawback of this method is that, in general, Bell inequalities are only a necessary condition for the existence of local and realistic description of the investigated quantum system. Only in a few cases, see for instance [3–5], it has been proved that some Bell inequalities are necessary and sufficient condition for the existence of local realism. All these cases deal with two [3] or more than two [4, 5] qubits and two local observables measured at each side of the Bell experiment.

In [6] a general approach to the problem has been presented. It is possible to find all relevant inequalities that have to be fulfilled by the probabilities obtained by the measurement of any number of local observables on the system consisting of an arbitrary number of quantum objects, each of which described by a Hilbert space of arbitrary dimesnion, so that it can be described in terms of local realism. However, the number of inequalities that have to be examined grows extremely fast with the dimension of the problem, i.e., number of local observables, quantum objects and dimension of Hilbert space describing given objects. This makes the method practically useless, as shown in [7, 8].

Recent research shows that a different approach is possible. In [9, 10] methods of numerical linear optimization has been successfully applied to two qubit correlations with up to ten local observables being measured at each side of the experiment [9] and for two N-dimensional objects ($2 \leq N \leq 16$) with two local observables at each side of the experiment [10, 11]. In this approach one does not find Bell inequalities but finds the conditions under which for the given quantum system and quantum observables measured on it there exists a local hidden variable model reproducing quantum results. Additionally, this method can be directly applied to the analysis of experimental data.

The paper [10] is a good example of how important it is to know necessary and sufficient conditions for the existence of local realism in the given case. For instance in [12], it was shown for two N-dimensional entangled systems that the Clauser-Horne-Shimony-Holt (CHSH) inequality [13] is maximally violated by the factor of $\sqrt{2}$. The reason for this is that CHSH inequality is not a sufficient condition for the existence of local realism for two entangled objects each described by a Hilbert space of the dimension greater than two. Indeed, the results of [10] show that violations of local realism increase with the dimension of the systems.

In this paper we prove analytically that the violation of local realism for two maximally entangled qutrits (objects described by a three dimensional Hilbert space) observed via two unbiased three input and three output beamsplitters [14] is stronger than for two maximally entangled qubits [10]. Earlier numerical computations advocated such a violation but rigorous analytical evidence has so far been lacking except for the trivial case of qubit. Thus it is anticipated that analytical proofs should exist for higher dimensional quantum systems. Our present work on qutrit therefore constitutes the first such attempt to confirm the previous numerical claim. Moreover, we also see that the extension from qubit to qutrit is clearly non-trivial. In fact, a comparison of our results with the separability condition for so called generalised Werner states [15] can shed a new light on the relation between local realism and separability of bipartite quantum systems.

We consider the Bell type experiment in which two spatially separated observers Alice and Bob measure two non-commuting observables A_1, A_2 for Alice and B_1, B_2 for Bob on the maximally entangled state $|\psi\rangle$ of two qutrits

$$|\psi\rangle = \frac{1}{\sqrt{3}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B + |2\rangle_A|2\rangle_B), \tag{1}$$

where $|k\rangle_A$ and $|k\rangle_B$ describe k-th basis state of the qutrit A and B respectively. Such a state can be prepared with pairs of photons with the aid of parametric down conversion (see [14]), in which case kets $|k\rangle_A$ and $|k\rangle_B$ denotes photons propagating to Alice and Bob in mode k.

Here we consider the special case in which both observers measure observables defined by 6-port (three input and three output ports) beam splitter. The extended theory of such devices can be found in [14]. Here we give only a brief description.

Unbiased 6-port beamsplitter, which is called tritter, [14] is a device with the following property: if one photon enters into any single input port (out of the 3), its chances of exit are equally split between all 3 output ports. One can always build tritter with the distinguishing trait that the elements of its unitary transition matrix, $\hat{T}$, are *solely* powers of the 3-rd root of unity $\alpha = \exp{(i2\pi/3)}$, namely $T_{kl} = \frac{1}{\sqrt{3}}\alpha^{(k-1)(l-1)}$. In front of i-th input port of the tritter we put a phase shifter that changes the phase of the incoming photon by ϕ_i. These three phase shifts, which we denote for convenience as a "vector" of phase shifts $\vec{\phi} = (\phi_1, \phi_2, \phi_3)$, are macroscopic local parameters that can be changed by the observer. Therefore, tritter together with the three phase shifters performs the unitary transformation $\hat{U}(\vec{\phi})$ with the entries $U_{kl} = T_{kl}\exp(i\phi_l)$.

Alice and Bob measure the following observables

$$\begin{aligned} A(\phi_i) &= \hat{U}(\vec{\phi_i})|0\rangle\langle 0|\hat{U}(\vec{\phi_i})^\dagger + \alpha\hat{U}(\vec{\phi_i})|1\rangle\langle 1|\hat{U}^\dagger(\vec{\phi_i}) + \alpha^2\hat{U}(\vec{\phi_i})|2\rangle\langle 2|\hat{U}^\dagger(\vec{\phi_i}) \\ B(\theta_j) &= \hat{U}(\vec{\theta_j})|0\rangle\langle 0|\hat{U}(\vec{\theta_j})^\dagger + \alpha\hat{U}(\vec{\theta_j})|1\rangle\langle 1|\hat{U}^\dagger(\vec{\theta_j}) + \alpha^2\hat{U}(\vec{\theta_j})|2\rangle\langle 2|\hat{U}^\dagger(\vec{\theta_j}), \end{aligned} \tag{2}$$

where $i, j = 1, 2$ and where, for instance, $\vec{\phi_i}$ denotes the vector of local phase shifts for Alice in the i-th experiment. Please notice that we ascribe complex numbers to the results of measurements, i.e., to the "click" of the l-th detector we ascribe the number α^l. The justification of such an assigment can be found in [14]. It results in a very symmetrical complex correlation function

$$\begin{aligned} E(\phi_i, \theta_j) &= \langle\psi|A(\phi_i)B(\theta_j)|\psi\rangle \\ &= \frac{1}{3}(\exp(\phi_i^1 - \phi_i^2 + \theta_j^1 - \theta_j^2) + \exp(\phi_i^2 - \phi_i^3 + \theta_j^2 - \theta_j^3) \\ &+ \exp(\phi_i^3 - \phi_i^1 + \theta_j^3 - \theta_j^1)), \end{aligned} \tag{3}$$

where, for instance, ϕ_i^1 denotes the first phase shift at Alice's side in the i-th experiment. This correlation function retains the information about the correlations observed in the experiment. In fact, according to quantum mechanics the whole information that is accessible in the experiment are probabilities of coincidence firings of the detectors. It can be easily verified through the knowledge of the correlation function $E(\phi_i, \theta_j)$ one is able to calculate the probabilities of these coincidence "clicks" and in this way obtain the whole information about the correlations observed in the system.

Following [10] We define the strength of violation of local realism as the minimal noise admixture F_{thr} to the state (1) below which the measured correlations cannot be described by local realism for the given observables. Therefore, we assume that Alice and Bob perform their measurements on the following mixed state ρ_F

$$\rho_F = (1-F)|\psi\rangle\langle\psi| + F\rho_{noise}, \tag{4}$$

where $0 \leq F \leq 1$ and where ρ_{noise} is a diagonal matrix with entries equal to 1/9. This matrix is a totally chaotic mixture (noise), which admits a local and realistic description. For $F = 0$ (pure maximally entangled state) local realistic description does not exist whereas for $F = 1$ (pure noise) it does. Therefore, there exists some threshold value of F, which we denote by F_{thr}, such that for every $F \leq F_{thr}$ local and realistic description does not exist. The bigger the value of F_{thr}, the stronger is the violation of local realism. The correlation function for the state (4) reads $E^F(\vec{\phi_i}, \vec{\theta_j}) = (1-F)E(\vec{\phi_i}, \vec{\theta_j})$.

Let us now assume that Alice measures two observables defined by the following sets of phase shifts $\vec{\phi_1} = (0, \pi/3, -\pi/3), \vec{\phi_2} = (0,0,0)$ whereas Bob measures two observables defined by the sets of phase shifts $\vec{\theta_1} = (0, \pi/6, -\pi/6), \vec{\theta_2} = (0, -\pi/6, \pi/6)$. From numerical computations it is known [10, 11] that these sets of phases gives the highest F_{thr}. Straightforward calculations give the following values of the correlations functions for each experiment: $E^F(\vec{\phi_1}, \vec{\theta_1}) = E^F(\vec{\phi_2}, \vec{\theta_2}) = Q_1 = \frac{2\sqrt{3}+1}{6} - i\frac{2-\sqrt{3}}{6}, E^F(\vec{\phi_1}, \vec{\theta_2}) = Q_1^*, E^F(\vec{\phi_2}, \vec{\theta_1}) = Q_2 = -\frac{1}{3}(1+2i)$. From these complex numbers we can create a 2×2 matrix $\hat{Q}$ with the entries $(\hat{Q})_{ij} = E^F(\vec{\phi_i}, \vec{\theta_j})$.

Local realism implies the following structure of the correlation function that is to reproduce the quantum correlation function defined above

$$E_{LHV}(\vec{\phi}_i,\vec{\theta}_j)=\int d\lambda\rho(\lambda)A(\vec{\phi}_i,\lambda)B(\vec{\theta}_j,\lambda), \tag{5}$$

where for trichotomic measurements $A(\vec{\phi}_i,\lambda)=\alpha^m$ and $B(\vec{\theta}_j,\lambda)=\alpha^n$, $(m,n=1,2,3)$. Three-valued functions $A(\vec{\phi}_i,\lambda), B(\vec{\theta}_j,\lambda)$ represent the values of local measurements predetermined by local hidden variables, denoted by λ, for the specified local settings. This expression is an average over a certain local hidden variable distribution $\rho(\lambda)$ of certain *factorisable* matrices, namely those with elements given by $H_\lambda^{ij}=A(\vec{\phi}_i,\lambda)B(\vec{\theta}_j,\lambda)$. The symbol λ may hide very many parameters. However, since the only possible values of $A(\vec{\phi}_i,\lambda)$ and $B(\vec{\theta}_j,\lambda)$ are $1,\alpha,\alpha^2$ there are only 9 *different* sequences of the values of $(A(\vec{\phi}_1,\lambda),A(\vec{\phi}_2,\lambda))$, and 9 *different* sequences of the values of $(B(\vec{\theta}_1,\lambda),B(\vec{\theta}_2,\lambda))$, and consequently they form only 81 matrices $\hat{H}_\lambda$.

Therefore the structure of local hidden variable model of $E_{LHV}(\vec{\phi}_i,\vec{\theta}_j)$ reduces to discrete probabilistic model involving the average of all the 81 matrices $\hat{H}_\lambda$. Therefore, we replace the parameter λ by index k ($k=1,2,\ldots,81$) to which we ascribe the matrix $\hat{H}_k$ with entries $H_k^{ij}=\alpha^{k_i+l_j}$ $(i,j=1,2)$, where $k_1=[(k-1)/9]-1, k_2=[(k-1)/3]-1, l_1=1, l_2=k$ (please notice that $\alpha^{-1}=\alpha^2$) and where $[x]$ denotes the integer part of the number x. It can be checked that only first 27 matrices are different, which means that it suffices to consider only them. With this notation the correlation function $E_{LHV}(\vec{\phi}_i,\vec{\theta}_j)$ acquires the following simple form

$$E_{LHV}(\vec{\phi}_i,\vec{\theta}_j)=\sum_{k=1}^{27}p_kH_k^{ij}, \tag{6}$$

with, of course, the probabilities satisfying $p_k\geq 0$ and $\sum_{k=1}^{27}p_k=1$. From $E_{LHV}(\vec{\phi}_i,\vec{\theta}_j)$ we build the matrix $\hat{E}_{LHV}$.

Quantum predictions in form of the matrix $\hat{Q}^F$ can be recovered by local hidden variables if and only if

$$\hat{Q}^F=\sum_{n=1}^{27}p_n\hat{H}_n. \tag{7}$$

Now, we want to find the minimal possible F for which it is still possible to recover matrix $\hat{Q}^F$ using the probability distribution p_n and factorizable matrices $\hat{H}_n$. For convenience we define a new parameter $V=1-F$. Then the minimal F refers to maximal V.

Theorem: The maximal V equals $V_{thr}=\frac{6\sqrt{3}-9}{2}$.

Proof: First we observe that matrix $\hat{Q}^V$ can be written in the following way

$$\hat{Q}^V=V\left[\frac{2\sqrt{3}+1}{6}-i\frac{2-\sqrt{3}}{6}\right]I+V\left[-\frac{2\sqrt{3}-1}{6}+i\frac{2+\sqrt{3}}{6}\right]\vec{n}\cdot\vec{\sigma}, \tag{8}$$

where $\vec{n}=(n_x,n_y,n_z)=\left(-\frac{1}{2},\frac{\sqrt{3}}{2},0\right)$ (please notice that $|\vec{n}|=1$). One observes that $\hat{Q}^V$ commutes with the matrix $\hat{\mathcal{U}}=\vec{n}\cdot\vec{\sigma}$, which has only two nonzero entries $\mathcal{U}_{12}=\alpha^2,\mathcal{U}_{21}=\alpha$ and is unitary and hermitian. Furthermore, $\hat{\mathcal{U}}$ preserves the structure of matrices $\hat{H}_n$ in the sense that for every $n=1,2,\ldots,27$, $\hat{\mathcal{U}}\hat{H}_n\hat{\mathcal{U}}=\hat{H}_m$ for some $m=1,2,\ldots,27$. This is one to one mapping. One can also find that some matrices $\hat{H}_n$ are invariants with respect to transformation $\hat{\mathcal{U}}$. For further considerations it is necessary to have the list of pairs (n,m)

$$\{(1,8),(2,10),(3,24),(4,17),(5,19),(7,26),(9,15),(11,13),(12,27),(14,22),(16,20),(23,25)\}. \tag{9}$$

By considering the i-th pair (n,m) in the above list, we can define new matrices, $\hat{G}_i=\hat{H}_n+\hat{H}_m$. Thus, $\hat{G}_1=\hat{H}_1+\hat{H}_8$, $\hat{G}_2=\hat{H}_2+\hat{H}_{10}$ and so forth. The remaining invariant matrices are $\hat{H}_6,\hat{H}_{18},\hat{H}_{21}$. For convenience, we label them as $\hat{G}_{13}$ to $\hat{G}_{15}$, viz $\hat{G}_{13}=\hat{H}_6,\hat{G}_{14}=\hat{H}_{18},\hat{G}_{15}=\hat{H}_{21}$.

Suppose that we have the optimal solution (the solution for which $V=V_{thr}$), i.e., we have the probability distribution p_n so that $\hat{Q}^{V_{thr}}=\sum_{n=1}^{27}p_n\hat{H}_n$. Acting on both sides of this equation with matrix $\hat{\mathcal{U}}$ we get another optimal solution with the same V_{thr} (matrix $\hat{\mathcal{U}}$ commutes with $\hat{Q}^{V_{thr}}$) but with the new probability distribution p'_k, which can be obtained from the previous one by swapping probabilities belonging to the same pair, for instance, $p'_{10}=p_2$ and so on. Therefore, due to the above property, we can assume *without loosing generality* that in the optimal solution the probabilities referring to the same pair are equal. Therefore, we have reduced the number of relevant probabilities from 27 to 15. One can observe that *every* matrix $\hat{G}_k$ can be expressed by matrices $\hat{G}_1,\hat{G}_{10},\hat{G}_{13}$ by multiplying them by α, α^2 and -1. For instance, $\hat{G}_6=\alpha\hat{G}_{10},\hat{G}_{11}=-\hat{G}_{13}$ etc. Three matrices are the same: $\hat{G}_5=\hat{G}_3,\hat{G}_7=\hat{G}_4,\hat{G}_{11}=\hat{G}_9$, which further reduces the number of relevant probabilities from 15 to 12.

Having in mind the above properties we can write the optimal solution in the new form

$$\hat{Q}^{V_{thr}} = \sum_{k\neq 5,7,11} w_k \hat{G}_k, \tag{10}$$

remembering that now the normalization condition for probabilities w_k reads

$$2(w_1 + w_2 + w_3 + w_4 + w_6 + w_8 + w_9 + w_{10} + w_{12}) + w_{13} + w_{14} + w_{15} = 1. \tag{11}$$

Due to the fact that all $\hat{G}_k$ can be expressed by $\hat{G}_1$, $\hat{G}_{10}$, $\hat{G}_{13}$, we have

$$\begin{aligned}\hat{Q}^{V_{thr}} &= (w_1 + \alpha w_8 + \alpha^2 w_{12})\hat{G}_1 + (w_{10} + \alpha w_6 + \alpha^2 w_2)\hat{G}_{10} \\ &+[(w_{13} - w_9) + \alpha(w_{14} - w_3) + \alpha^2(w_{15} - w_4)]\hat{G}_{13}.\end{aligned} \tag{12}$$

Notice that $\hat{G}_1 + \hat{G}_{10} - \hat{G}_{13} = 0$ so that we have

$$\begin{aligned}\hat{Q}^{V_{thr}} &= [(w_1 + w_{13} - w_9) + \alpha(w_8 + w_{14} - w_3) + \alpha^2(w_{12} + w_{15} - w_4)]\hat{G}_1 \\ &+[(w_{10} + w_{13} - w_9) + \alpha(w_6 + w_{14} - w_3) + \alpha^2(w_2 + w_{15} - w_4)]\hat{G}_{10}.\end{aligned} \tag{13}$$

Matrix $\hat{G}_1$ (with entries $G_1^{11} = 2$, $G_1^{12} = -\alpha^2$, $G_1^{21} = -\alpha$, $G_1^{22} = 2$) is a sum of matrices $\hat{H}_1, \hat{H}_8$ whereas matrix $\hat{G}_{10}$ (with entries $G_{10}^{11} = -1$, $G_{10}^{12} = 2\alpha^2$, $G_{10}^{21} = 2\alpha$, $G_{10}^{22} = -1$) is a sum of matrices $\hat{H}_{14}, \hat{H}_{22}$. These four matrices are linearly independent so they form a basis in four dimensional space of 2×2 complex matrices. The expansion of $\hat{Q}^{V_{thr}}$ in this basis reads

$$\hat{Q}^{V_{thr}} = \lambda_1 \hat{G}_1 + \lambda_{10}\hat{G}_{10}, \tag{14}$$

where $\lambda_1 = V_{thr}[(\frac{1}{6}+\frac{1}{3\sqrt{3}})+i(-\frac{1}{9}+\frac{1}{2\sqrt{3}})], \lambda_{10} = V_{thr}[(\frac{1}{6}-\frac{1}{3\sqrt{3}})+i(\frac{1}{9}+\frac{1}{2\sqrt{3}})]$. Because both λ_1 and λ_{10} lie on the complex plane between complex numbers 1 and α, they can be uniquely expressed by these numbers 1 and α with positive coefficients, i.e., $\lambda_1 = V_{thr}[\frac{1}{27}(9+2\sqrt{3}) + \alpha\ \frac{1}{27}(9-2\sqrt{3})]$ and $\lambda_{10} = V_{thr}[\frac{1}{27}(9-2\sqrt{3}) + \alpha\ \frac{1}{27}(9+2\sqrt{3})]$.

We can rewrite the formula (13) using the indentity $1+\alpha+\alpha^2 = 0$ in the following form

$$\begin{aligned}\lambda_1 &= (w_1 + w_4 + w_{13} - w_9 - w_{12} - w_{15}) + \alpha(w_4 + w_8 + w_{14} - w_3 - w_{12} - w_{15}), \\ \lambda_{10} &= (w_4 + w_{10} + w_{13} - w_2 - w_9 - w_{15}) + \alpha(w_4 + w_6 + w_{14} - w_2 - w_3 - w_{15}).\end{aligned} \tag{15}$$

After comparing Eq.(15) and Eq.(14) we have

$$\begin{aligned}w_1 + w_4 + w_{13} - w_9 - w_{12} - w_{15} &= \frac{V_{thr}}{27}(9+2\sqrt{3}), \\ w_4 + w_8 + w_{14} - w_3 - w_{12} - w_{15} &= \frac{V_{thr}}{27}(9-2\sqrt{3}), \\ w_4 + w_{10} + w_{13} - w_2 - w_9 - w_{15} &= \frac{V_{thr}}{27}(9-2\sqrt{3}), \\ w_4 + w_6 + w_{14} - w_2 - w_3 - w_{15} &= \frac{V_{thr}}{27}(9+2\sqrt{3})\end{aligned} \tag{16}$$

Because we deal with the optimal solution for which V is maximal $V = V_{thr}$ all the probabilities with negative sign in (16) must be zero (please notice that none of the probabilities that come into (16) with negative sign appears in any equation with a positive sign). We get

$$\begin{aligned}w_1 + w_4 + w_{13} &= \frac{V_{thr}}{27}(9+2\sqrt{3}), \\ w_4 + w_8 + w_{14} &= \frac{V_{thr}}{27}(9-2\sqrt{3}), \\ w_4 + w_{10} + w_{13} &= \frac{V_{thr}}{27}(9-2\sqrt{3}), \\ w_4 + w_6 + w_{14} &= \frac{V_{thr}}{27}(9+2\sqrt{3}).\end{aligned} \tag{17}$$

Now the whole probability distribution consists of $w_1, w_4, w_6, w_8, w_{10}, w_{13}, w_{14}$. By substracting the fourth equation from the second one and the third one from the first one we arrive at

$$w_6 - w_8 = \frac{4\sqrt{3}}{27} V_{thr}, \quad w_1 - w_{10} = \frac{4\sqrt{3}}{27} V_{thr}. \tag{18}$$

Again, because V_{thr} is maximal, it must be $w_8 = w_{10} = 0$. Thus $w_1 = w_6 = \frac{4\sqrt{3}}{27} V_{thr}$ and the second and the third equation in (16) become

$$w_4 + w_{14} = \frac{V_{thr}}{27}(9 - 2\sqrt{3}), \quad w_4 + w_{13} = \frac{V_{thr}}{27}(9 - 2\sqrt{3}), \tag{19}$$

This clearly implies $w_{13} = w_{14} = q$. Normalization condition now reads

$$2(w_1 + w_4 + w_6) + w_{13} + w_{14} = 1. \tag{20}$$

A simple algebra gives

$$q + w_4 = \frac{1}{2} - \frac{8\sqrt{3}}{27} V_{thr}. \tag{21}$$

However, from (19), we know that $q + w_4 = \frac{V_{thr}}{27}(9 - 2\sqrt{3})$. Therefore

$$\frac{1}{2} - \frac{8\sqrt{3}}{27} V_{thr} = \frac{V_{thr}}{27}(9 - 2\sqrt{3}) \tag{22}$$

which gives $V_{thr} = \frac{6\sqrt{3}-9}{2}$. This ends the proof.

We have shown analytically that for the Bell experiment with the four trichotomic observables (2) (two at each side of the experiment) defined by the sets of phase shifts $\vec{\phi}_1 = (0, \pi/3, -\pi/3)$, $\vec{\phi}_2 = (0, 0, 0)$, $\vec{\theta}_1 = (0, \pi/6, -\pi/6)$, $\vec{\theta}_2 = (0, -\pi/6, \pi/6)$ the minimal noise admixture F_{thr} above which local and realistic description exists is $F_{thr} = 1 - V_{thr} = \frac{11-6\sqrt{3}}{2}$. For two maximally entangled qubits this number is $\frac{2-\sqrt{2}}{2} < F_{thr}$. Therefore, two entangled qutrits are more robust against local and realistic description than two entangled qubits.

Although, the presented here proof cannot be easily applied to the set of arbitrary observables defined in (2) as it relies on the symmetry properties of matrix $\hat{Q}^F$ it may be considered as the first step towards the Bell theorem for two entangled qutrits.

MZ thanks Anton Zeilinger and Alipasha Vaziri for discussions. MZ and DK are supported by the University of Gdansk Grant No. BW/5400-5-0032-0 and by KBN grant No. 5 P03B 088 20. This paper is also supported in part under NUS research grant No. R-144-000-054-112.

---

[1] J. Bell, Physics **1**, 195 (1964).
[2] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, Phys. Rev. Lett. **23**, 15, 880 (1969); E. P. Wigner, Am. J. Phys. **38**, 8, 1005 (1970); J. F. Clauser and M. A. Horne, Phys. Rev. D **10**, 526 (1974); N. D. Mermin, Phys. Rev. D **22**, 2, 356 (1980); S. L. Braunstein and C. M. Caves, Ann. Phys. (NY) **202**, 22 (1990). N. D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990); M. Ardehali, Phys. Rev. D. **44**, 10, 3336 (1991); N. Gisin and A. Peres, Phys. Lett. A. **162**,1,15 (1992); A. V. Belinskii, D. N. Klyshko, Phys. Usp. **36**, 653 (1993); M. Żukowski, Phys. Lett. A **177**,4-5,290 (1993); N. Gisin, H. Bechmann-Pasquinucci, Phys. Lett. A **246**, 1 (1998); N. Gisin, Phys. Lett. A **260**, 1 (1999); D. Kaszlikowski and M. Żukowski, Phys. Rev. A **61**,2, 022114 (2000); I. Pitovsky, K. Svozil, quant-ph//0011060.
[3] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[4] R. F. Werner, M. M. Wolf, quant-ph//0102024.
[5] M. Żukowski and C. Brukner, quant-ph//0102039.
[6] N. D. Mermin and G. Schwarz, Found. Phys. **12**,101 (1982).
[7] I. Pitovsky, Math. Programming **50**, 395 (1991).
[8] A. Peres, Found. Phys. **29**, 589 (1999).
[9] M. Żukowski, D. Kaszlikowski, A. Baturo, J-A. Larsson, quant-ph//9910058.
[10] D. Kaszlikowski, P. Gnaciński, M. Żukowski, W. Miklaszewski and A. Zeilinger, Phys. Rev. Lett. **85**, 4418 (2000).
[11] T. Durt, D. Kaszlikowski and M. Żukowski, quant-ph//0101084.
[12] N. Gisin and A. Peres, Phys. Lett. A **162**, 15 (1992).
[13] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[14] M. Żukowski, A. Zeilinger, M. A. Horne, Phys. Rev. A **55**, 2564 (1997).
[15] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).